Spotlight on Four-steps SRL decomposition

Alessandro Moschitti¹, Ana-Maria Giuglea¹, Bonaventura Coppola², Roberto Basili¹

¹University of Rome “Tor Vergata”
²ITC-Irst and University of Trento

CoNLL-2005 Shared Task
Two-steps SRL decomposition

- Many systems exploit a traditional decomposition of the SRL task into Argument Identification and Argument Classification (i+c)
- In the Shared Task 2005, 11 systems out of 19 exploit an i+c strategy (Carreras and Màrquez, CoNLL 2005)
- Underlying assumption: verb arguments (including adjuncts) share common properties which are not dependent on their semantic role.
- Benefit: reduced computational effort in Argument Classification (one versus all)
Four-steps SRL decomposition

- Both Identification and Classification are further split

High-recall binary SVM classifier

Tree-kernel-based PAST classifier (Moschitti et al., ACL05 Workshop on Feature Engineering)

Two-stages Hierarchical Argument Classifier

52 Individual PropBank Role Classifiers
Impact of Hierarchical Classification

- Simple and hierarchical role classifiers compared over gold boundaries. Development set (sec. 24) used for testing. F1 reported.

<table>
<thead>
<tr>
<th></th>
<th>AX</th>
<th>AM</th>
<th>CX</th>
<th>RX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train Instances</td>
<td>172457</td>
<td>59473</td>
<td>2954</td>
<td>7928</td>
</tr>
<tr>
<td>Test Instances</td>
<td>5930</td>
<td>2132</td>
<td>105</td>
<td>284</td>
</tr>
<tr>
<td>Step 3: binary class.</td>
<td>97.29</td>
<td>97.35</td>
<td>70.86</td>
<td>93.15</td>
</tr>
<tr>
<td>Step 3: combination</td>
<td></td>
<td></td>
<td></td>
<td>95.99</td>
</tr>
<tr>
<td>Step 4 multi-c (over gold 3)</td>
<td>92.50</td>
<td>85.88</td>
<td>91.43</td>
<td>91.55</td>
</tr>
<tr>
<td>Steps 3+4 comb</td>
<td></td>
<td></td>
<td></td>
<td>88.15</td>
</tr>
<tr>
<td>Basic 2-steps system</td>
<td></td>
<td></td>
<td></td>
<td>88.61</td>
</tr>
</tbody>
</table>
Discussion of Results

• Introducing an intermediate classification layer (step 3) doesn’t noticeably reduce performance

• Hierarchical Classification allows for independent tuning of AXs, AMs, RXs, and CXs classifiers with respect to:
  • Parameter optimization
  • Feature selection

• Benefit: Further reduction of processing time

• Future Work: Possible exploitation of Tree Kernels
Questions for Public Discussion

• Did you find/analyze different behaviors of argument classes with respect to individual features?

• Did you attempt any per-class feature selection?

• Can you report on previous experience in splitting multi-classification problems into several layers?